

SelectFeaturesColumns Documentation

Description: Takes a 'column slice' from a .res, .gct, .snp, .cn, .odf, or .cls file. **Author:** Joshua Gould (Broad Institute), gp-help@broad.mit.edu

Summary: This module generates a new file based on features (columns) selected from an existing file. For example, a res file can be produced containing only the tumor samples from a larger res file containing data for both tumor and normal samples. Specific rows can be extracted using the *SelectFeaturesRows* module.

Use the *columns* parameter to select the features to extract into the new file. Features are numbered beginning at zero. When entering a value for the *columns* parameter, you are choosing the features to extract, not the data columns to extract. An input file may contain one or more columns per feature. For example, if your input file is a .res file that contains three samples (features), you might enter *columns* =1 to extract the second sample:

Input res file:

Description	Accession	sample1		sample2		sample3	
	sample descriptions (which are ignored)						
5							
gene1 description	gene1_name	-214	Α	-135	Α	-106	Α
gene2 description	gene2_name	-153	Α	-114	Α	-125	Α
gene3 description	gene3_name	-58	Α	265	Α	-76	Α
gene4 description	gene4_name	88	Α	12	Α	168	Α
gene5 description	gene5_name	-295	Α	-419	Α	-230	Α

Output res file for columns =1:

Catpat res lie for columns 1:							
Description	Accession	sample2					
	sample descriptions						
5							
gene1 description	gene1_name	-135	Α				
gene2 description	gene2_name	-114	Α				
gene3 description	gene3_name	265	Α				
gene4 description	gene4_name	12	Α				
gene5 description	gene5_name	-419	Α				

Parameters:

Name	Description
input filename	input fileres, .gct, .snp (non-allele-specific), .cn, .odf, or .cls file
	from which the slice is extracted
columns	columns to select (e.g. 0-3, 5, 9). Either columns or a list file must
	be provided
list filename	text file containing one identifier per line that specifies the
	columns to select. Either columns or a list file must be provided.
output	output filenameres, .gct, .snp, .cn, .odf, .cls

GenePattern

Output Files:

1. output file containing only the specified features/samples (columns)

Platform Dependencies:

GeneListSelection

Module type: CPU type: any OS: any Language: Java